

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: November 20, 2002, 14:47:28 ; Search time 355 Seconds

(without alignments)  
3391.245 Million cell updates/sec

Title: US-09-802-285-2

Perfect score: 2043  
Sequence: 1 MKKQILYLVQQLFLCSAY.....VGNSTVPYTNLSCSETAR 384

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 2185239 segs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ .p2n.model -DEV=xih  
-O=/cgn2\_1/USPTO.spool/US09802285/runat.15112002.114655.18334/app-query.fasta.1.583  
-DB=N-Geneseg.101002 -OFMT=fastap -SUFFIX=ring -XINMATCH=0.1 -LOOPL=0  
-LOOPEXT=0 -UNITS=bis -START=1 -END=1 -MATRIX=blonsum62 -TRANS=human40.cdi  
-LIST=45 -DOCLIG=200 -THR.SCORE=pcr -THR.MAX=100 -THR.MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=plc -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USRR=US09802285 -OCGN.1.1.125 -runat.15112002.114655.18334 -NCPU=6 -ICPU=3  
-NO.XLPY -NO.MAP -LARGEDEVRY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-WARN.TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

N\_Geneseg.101002:\*  
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*  
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*  
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*  
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2043	100.0	1379	18 AAT87488	Mutant polysacchar
2	2043	100.0	1379	21 AAT51243	F. heparinum hepar
3	2043	100.0	1379	23 ABLS0563	Flavobacterium hep
4	2040	99.9	1379	14 AAQ41046	Heparinase. Flavo
5	116	5.7	2538	24 ABN70382	Streptococcus poly
6	116	5.7	2562	24 ABN67361	Streptococcus poly
7	111	5.4	2365589	24 ABA90521	Genomic sequence o
8	107	5.2	2178	20 AA212261	Neisseria gonorrhoe
9	105	5.1	2178	20 AA212259	Neisseria meningit
10	105	5.1	2178	20 AA212260	Neisseria meningit
11	105	5.1	47475	21 AA481465	N. meningitidis pa
12	105	5.1	349980	21 AAF21612	Neisseria meningit
13	105	5.1	837096	21 AA481489	N. meningitidis pa
14	103	5.0	5069	22 AAK52292	Human polynucleoti
15	103	5.0	5151	23 ABV22238	Human prostate exp
16	103	5.0	5151	23 ABV28076	Human prostate exp
17	103	5.0	5465	22 AAS60778	Human cancer agent
18	102.5	5.0	6300	21 ABA51864	B. subtilis opp op
19	102	5.0	1719	24 ABO68105	Listeria monocytog
20	102	5.0	1719	24 ABO69946	Listeria monocytog
21	101.5	5.0	1146	22 AAH01425	Klebsiella pneumon
22	101.5	5.0	6135	24 ABO69931	Listeria monocytog
23	101	4.9	1707	20 AAX20032	Enterococcus faeca
24	101	4.9	10955	20 AABN8017	E faecalis EF018 g
25	101	4.9	10955	20 AAX13173	Enterococcus faeca
26	98	4.8	1653	24 ABBN67396	Streptococcus poly
27	97.5	4.8	2277	24 ABR74276	Bacillus lichenillo
28	97	4.7	3721	20 AAX13348	Enterococcus faeca
29	96	4.7	2070	23 ABL10733	Drosophila melanog
30	95	4.7	1461	22 AAH44660	Equine infectious
31	95	4.7	4768	21 AACT79604	Virulence gene #24
32	95	4.7	8255	22 AAH44659	Equine infectious
33	95	4.7	213251	22 AAB67193	Listeria innocua c
34	94.5	4.6	1540	20 AAX20033	Enterococcus faeca
35	94.5	4.6	1540	24 AABN8018	E faecalis EF018 g
36	94.5	4.6	29069	21 AAB81497	N. meningitidis pa
37	94.5	4.6	349980	21 AAF21609	Neisseria meningit
38	94.5	4.6	1437668	21 AAB81490	N. meningitidis B
39	94	4.6	1664976	19 AAV21209	Methanococcus jam
40	93.5	4.6	4557	24 ABR75361	Bacillus lichenillo
41	93	4.6	1149	21 AAB44838	Arabidopsis thalia
42	93	4.6	2313	24 ABO69770	Listeria innocua D
43	93	4.6	3108	21 AAB79734	Eucalyptus grandis
44	93	4.6	684707	24 ABO67196	Listeria innocua c
45	92.5	4.5	1580	19 AAV19363	Cyclomaltoextrin

ALIGNMENTS

RESULT 1	
AAT87488	
ID AAT87488 standard; DNA: 1379 BP.	
XX	
AC AAT87488;	
XX	
DT 02-FEB-1998 (first entry)	
XX	
DE Mutant polysaccharide lyase derived from F. heparinum heparinase.	
XX	
KM Polysaccharide: lyase; Flavobacterium heparinum; heparinase; mutant;	
KW heparine; ds.	
XX	
OS Flavobacterium heparinum.	
XX	
Key Location/Qualifiers	
FT 173..1327	
FT /*tag= a	

FT	mat_peptide	236..1324
FT		/**tag- b
XX		
XX	M09716556-A1.	
PD	09-MAY-1997.	
XX		
PF	30-OCT-1996:	96MO-U017310.
PR	30-OCT-1995:	95U0-0008069.
PA	(MASI ) MASSACHUSETTS INST TECHNOLOGY.	
PI	Cooney CL, Ernst S, Godavarti R, Langer R, Sasisekharan R;	
PI	Venkataraman G;	
DR	WP1: 1997-272124/24.	
DR	P-PSDB: AAM28545.	
PT	New polysaccharide lyase derived from Flavobacterium heparinum	
PT	heparinase - by specified mutation(s); used to remove heparin from	
PT	blood and to digest heparin to low molecular weight fragments	
XX		
PS	Claim 15; Page 53-57; 73pp; English.	
XX		
CC	The new polysaccharide lyase has the amino acid sequence encoded	
CC	by the sequence given in AAT87488 with at least one of the	
CC	following amino acid substitutions:	
CC	(a) Cys135 replaced by Asp, Glu, Ser, Thr or His,	
CC	(b) a conservative substitution in the Cardin-Weintaub-like	
CC	heparin-binding sequence (HBS) XBKXKBXB (positions 197-205	
CC	or 208-212; B= any basic residue, X= any hydrophobic or	
CC	other residue) with a residue that conforms to the HBS,	
CC	(c) a conservative substitution in a EF-hand-like calcium	
CC	binding site (CBS), position 206-220, with a residue that	
CC	conforms to CBS,	
CC	(d) a (non-)conservative substitution of a PB1, PB2 or PB3 beta-	
CC	sheet domain, provided this preserves the parallel beta-	
CC	helix structure,	
CC	(e) non-conservative substitution of Cys297,	
CC	(f) deletion of one or more N- or C-terminal residues, provided	
CC	the parallel beta-helix structure is retained,	
CC	(g) substitution of His203 by Asp, Glu, Ser, Thr or Cys,	
CC	(h) substitution of Lys198, 199, 203, 208, 208, 211 or 214 with	
CC	a small (non-)polar or acidic amino acid,	
CC	(i) substitution of Phe197, Asn200, Asp204, Gu207, Asp210,	
CC	Asp212 or Gly213 by a small (non-) polar amino acid,	
CC	(j) non-conservative substitution of Ser39.	
CC	The mutant lyase is used to remove heparin from fluids, esp.	
CC	circulating blood. The lyase may have (i) altered enzymatic	
CC	activity compared with the native enzyme, or (ii) same activity	
CC	but reduced immunogenicity (particularly where Ser39 is replaced).	
CC	Also they may be made more stable and easier to produce and purify	
CC	(pref. by removal of crosslinkable Cys residues).	
XX		
SQ	Sequence 1379 BP; 465 A; 308 C; 281 G; 325 T; 0 other;	
Alignment Scores:		
Pred. NO.:	6.4e-213	Length: 1379
Score:	2043.00	Matches: 384
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	18	Gaps: 0

OY 1 MetLysLVSGInIleLeuTyrIreulleValLeuGInGInLeuPheLeuLysSerLaIayr 20  
Db 173 ATGAATAAACCAATTCTATATCTGATTGACTTCAGCAACTGTTCCTCGTTCCGGCTTAC 23  
Oy 21 AlaGInGInLysSerGlyAsnIleProTyrArgValAsnValGlnAlaSpSerAla 40

Db	233	GCCAGCAAAAAAATCCGGTAAACATCCCTTACCGGGTAAATGTGCAGGCGACAGTCT	292
Qy	41	LysGlnLysAlaIleIleAspAsnLysTrpValAlaValGlyIleAsnLysProTyrAla	60
Db	293	AAGCAGAAAGCGATTATTGTGCAACAATATGGGGGAGTAGAGCATCAATAAACCTTATGCA	352
Qy	61	LeuGlnTyrAspAspLysLeuArgPheAsnGlyLysProSerTyrArgPheGluLeuLys	80
Db	353	TTNCAATATACCATTAACACTGGCCTTATATGAAAAACCTCTATCGCTTGGAGCTTAA	412
Qy	81	AlaGluAspAsnSerLeuGluGlyTyrAlaAlaGlyGluThrLysGlyArgThrGluLeu	100
Db	413	GCCAAACAAATTCGCTTGAAGGTTATCGTCGAGGAGAAACAAAGGCGGTACAAATTG	472
Qy	101	SerTyrSerTyrIleThrThrAsnAspPheLysLysPheProProSerValTyrGlnAsn	120
Db	473	TCGTACAGCATGCAACACCAATGATTTTAAAGAAATTTCCCAAGGGTATACCAAAAT	532
Qy	121	AlaGlnLysLeuLysThrValTyrHisTyrGlyLysGlyIleCysGluGlnGlySerSer	140
Db	533	GCGCAAAAGCTAAAAACCGTTTATCATATACGGAAAGGATTTGTGAACAGGAGGCTCC	592
Qy	141	ArgSerTyrThrPheSerValTyrIleProSerSerPheProAspAsnAlaThrIle	160
Db	593	CGCAGCATACCTTTTCAAGTACATACCTCCCTCCCGCAAAATCGCATCTATT	652
Qy	161	PheAlaGlnTrpHisGlyAlaProSerArgThrLeuValAlaThrProGluGlyIle	180
Db	653	TTTGCCCAATGGCATGGGTGCACCCACAGACAGCTGTGTACATACACAGAGGACAAATT	712
Qy	181	LysThrLeuSerIleGluGluPheLeuAlaLeuTyrAspArgMetIlePheLysAsn	200
Db	713	AAAACACTGAGCATAGAGAGTTTGGCTTATACAGCCGATATCTTCAAAAAAAT	772
Qy	201	IleAlaHisAspLysValGluLysLysAspLysAspGlyLysIleThrTyrValAlaGly	220
Db	773	ATGCCCATGTATAAGTTGAAAAAAGATTAAGACGAAAAATTACTTATGTACCGGA	832
Qy	221	LysProAsnGlyTyrLysValGluGlnGlyGlyTyrProThrLeuAlaPheGlyPheSer	240
Db	833	AAGCCAATGGCTGGAAGTAGAAGAACAAAGGGTTATCCACAGCTGGCTTGGTTTCT	892
Qy	241	LysGlyLysThrPheTyrIleLysAlaAsnSerAspArgGlnTrpLeuThrAspLysAlaSp	260
Db	893	AAAGGATTTTTTATACATCAAGGCAAACTCGACGGGAGTGGCTTACCAACAAACCGAC	952
Qy	261	ArgAsnAsnAlaAsnProGluAsnSerGluValMetLysProTyrSerSerGluTyrLys	280
Db	953	CGTAAACATGCCAATCCCGAGATTAAGTAAGTAATAACCCCTATTCTCGGAATACAA	1012
Qy	281	ThrSerThrIleAlaTyrLysMetProPheAlaGlnPheProLysAspCysTrpIleThr	300
Db	1013	ACTTCAACCATTTGGCTTAAATAATGCCCTTGGCCCACTTCCATAAGATGTGCGATTACT	1072
Qy	301	PheAspValAlaIleAspTrpThrLysTyrGlyLysGluAlaIleAsnThrIleLeuLysPro	320
Db	1073	TTTATGTCCGCATAGCTGACGACGAATATGAAAAAGAGCCAAATTTTGAAACCC	1132
Qy	321	GlyLysLeuAspValMetMetThrTyrThrLysAsnLysLysProGlnLysAlaHisIle	340
Db	1133	GGTAGCTGTGATGATGATGAGACTTATACCAAAATAAACAACCAAAAACCGCATATC	1192
Qy	341	ValAsnGlnGlnGluIleLeuIleGlyArgAsnAspAspAspGlyTyrTyrPheLysPhe	360
Db	1193	GTAACAACGACGAGAAATCCGATCGGACGTAACGAATGACGATTAATCTCAAAATT	1252
Qy	361	GlyIleTyrArgValGlyAsnSerThrValProValThrTyrAsnLeuSerGlyTyrSer	380
Db	1253	GAATATTACAGGGTCGGTATACACACAGGTCGCCGGTACTTATTAACCTGACGGCGTACAC	1312
Qy	381	GluThrAlaArg 384	
Db	1313	GAAACTGCCAGA 1324	

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame-plus\_p2n model

Run on: November 20, 2002, 14:54:03 ; Search time 52 Seconds  
(Without alignments)  
2264.690 Million cell updates/sec

Title: US-09-802-285-2

Sequence: 1 MKKQILYLVQLQFLCSAY.....VGNSTVPVYVNLGSETAR 364

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 88:724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+<sub>p2n</sub>.model -DEV=xlh  
-O=/cgn2\_1/USPTO/US09802285/runat\_15112002\_11657\_18362/app\_query.fasta\_1.583  
-DB=Issued\_Patents\_NA -QEMT=fastap -SUFFIX=trn1 -MINMATCH=0.1 -IOOCL=0  
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blotsum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR\_SCORE=pcr -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pcr -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09802285 @CGN1.1.1.1 @runat\_15112002\_11657\_18362 -MCPU=6 -ICPU=3  
-NO\_XLIPX -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELDP=6 -DELEX=7

Database :

- Issued\_Patents\_NA:\*
- 1: /cgn2\_6/prodata/1/ina/5A.COMB.seq:\*
  - 2: /cgn2\_6/prodata/1/ina/5B.COMB.seq:\*
  - 3: /cgn2\_6/prodata/1/ina/6A.COMB.seq:\*
  - 4: /cgn2\_6/prodata/1/ina/6B.COMB.seq:\*
  - 5: /cgn2\_6/prodata/1/ina/PCTUS.COMB.seq:\*
  - 6: /cgn2\_6/prodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2043	100.0	1379	1	US-07-783-706-1 / 7020
2	2043	100.0	1379	5	PCT-US92-09124-1
3	2043	100.0	1407	4	US-09-066-481-1
4	2038	99.8	1379	2	US-08-445-342A-1
5	101	4.9	1707	4	US-09-071-035-65
6	94.5	4.6	1540	4	US-09-071-035-67
7	92.5	4.5	3695	4	US-09-071-035-419
8	92.5	4.5	3840	4	US-09-071-035-417
9	92	4.5	11864	4	US-08-961-527-61
10	91	4.5	4661	4	US-08-961-527-61
11	90.5	4.4	2630	3	US-08-666-286-6
12	90.5	4.4	2630	4	US-09-469-253-6

13	90.5	4.4	2630	4	US-09-642-146-6	Sequence 6, Appli
14	90	4.4	980	4	US-09-221-017B-589	Sequence 589, App
15	90	4.4	2860	4	US-09-221-017B-871	Sequence 871, App
16	90	4.4	6885	1	US-08-570-311-9	Sequence 9, Appli
17	90	4.4	6895	2	US-08-353-485-9	Sequence 1, Appli
18	89.5	4.4	3827	2	US-08-447-031A-1	Sequence 1, Appli
19	89.5	4.4	4612	2	US-08-447-031A-8	Sequence 8, Appli
20	89	4.4	2191	4	US-09-071-035-363	Sequence 363, App
21	89	4.4	2334	4	US-09-071-035-361	Sequence 361, App
22	89	4.4	4310	3	US-09-008-172-1	Sequence 1, Appli
23	89	4.4	4310	4	US-09-210-361-5	Sequence 5, Appli
24	89	4.4	4310	4	US-09-740-274-5	Sequence 5, Appli
25	88	4.3	4460	3	US-09-007-999-1	Sequence 1, Appli
26	88	4.3	4460	4	US-09-210-361-1	Sequence 1, Appli
27	88	4.3	4460	4	US-09-740-274-1	Sequence 7, Appli
28	87.5	4.3	1169	4	US-08-386-642-7	Sequence 9, Appli
29	87.5	4.3	2061	1	US-08-204-656B-9	Sequence 5, Appli
30	87.5	4.3	2061	1	US-08-470-702-5	Sequence 5, Appli
31	87.5	4.3	2061	1	US-08-467-831-5	Sequence 4, Appli
32	87.5	4.3	4344	2	US-08-532-547-4	Sequence 4, Appli
33	87.5	4.3	4344	2	US-08-379-656B-4	Sequence 4, Appli
34	87.5	4.3	4344	3	US-08-455-838-4	Sequence 4, Appli
35	87.5	4.3	4344	3	US-09-019-809-4	Sequence 4, Appli
36	87.5	4.3	4344	4	US-09-471-177-4	Sequence 4, Appli
37	87.5	4.3	9711	4	US-08-961-527-167	Sequence 167, App
38	87	4.3	2973	1	US-08-451-715A-7	Sequence 7, Appli
39	87	4.3	3533	2	US-08-476-062A-40	Sequence 40, Appli
40	87	4.3	3533	6	PCT-US96-01314-40	Sequence 40, Appli
41	87	4.3	3533	6	5424399-1	Patent No. 5424399
42	87	4.3	6836	4	US-08-976-259-73	Sequence 73, Appli
43	86.5	4.2	1585	4	US-09-071-035-43	Sequence 43, Appli
44	86.5	4.2	1659	4	US-09-071-035-41	Sequence 41, Appli
45	86.5	4.2	1983	4	US-09-134-001C-1524	Sequence 1524, Ap

# ALIGNMENTS

PD Feb 3 '98

RESULT 1  
US-07-783-706-1  
Sequence 1, Application US/07783706  
Patent No. 5714376  
GENERAL INFORMATION:  
APPLICANT: Sasisekharan, Ramnath  
APPLICANT: Moreman, Kelley L.  
APPLICANT: Cooney, Robert S.  
TITLE OF INVENTION: The Hepatrinase Gene from Flavobacterium  
TITLE OF INVENTION: Hepatrinum  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kilpatrick & Cody  
STREET: 100 Peachtree Street, Suite 3100  
CITY: Atlanta  
STATE: Georgia  
COUNTRY: US  
ZIP: 30303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/783,706  
FILING DATE: 19911023  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: MIT5546  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 404-572-6508  
TELEFAX: 404-572-6555  
INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:  
LENGTH: 1379 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHEICAL: YES  
ANTI-SENSE: YES  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
ORGANISM: Flavobacterium heparinum  
US-07-783-706-1

Alignment Scores:  
Pred. No.: 1.34e-249 Length: 1379  
Score: 2043.00 Matches: 384  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 1 Gaps: 0  
US-09-802-285-2 (1-384) x US-07-783-706-1 (1-1379)

QY 1 MetLysLysGlnIleLeuTyrLeuIleValIleuGlnGlnLeuPheLeuCysSerAlaTyr 20  
Db 173 ATGAAAAACAATTCATCTATCTGATCTTCTGCAACTGCTTCTTGTGGCTTAC 232  
QY 21 AlacInLysLysSerGlyAsnIleProTyrArgValAsnValGlnAlaAspSerAla 40  
Db 233 GCCCAGCAAAAAAATCCGGTAACATCCCTTACCGGGTAATGTGACAGCCGACAGTCT 292  
QY 41 LysGlnLysAlaIleIleAspAsnLysTyrValAlaValGlyIleAsnLysProTyrAla 60  
Db 293 AAGCAGAGAGGATATTATGACACAAATGGGTGGCAGTACATCAATTAACCTATGCA 352  
QY 61 LeuGlnTyrAspAspLysLeuArgPheAsnGlyLysProSerTyrArgPheGlnLeuLys 80  
Db 353 TTACAATATGCGATTAACCTCGCTTTATATGAAAAATCCATCCCTTGTGAGCTTAAA 412  
QY 81 AlacLysAspAsnSerLeuGlnGlyTyrAlaAlaGlyLysGlnTyrLysGlyArgThrGlnLeu 100  
Db 413 GCCCAACACATTCGCTTGAAGCTTATGCTGACAGAGAAACAAAGGCGCTACAGAAATG 472  
QY 101 SerTyrSerTyrAlaThrThrAspAspPheLysLysPheProPheProSerValTyrGlnAsn 120  
Db 473 TCGTACAGCTATGCAACACCAATGATTTTAAGAAATTTCCCAAGCGTATACCAAAAT 532  
QY 121 AlacInLysLeuLysThrValTyrHisTyrGlyLysGlyIleCysGlnLysGlySerSer 140  
Db 533 GCGCAAAAGCTAAACCCGTTATCATCTACGCAAAAGGATTTGTGAACAGGAGAGCTCC 592  
QY 141 ArgSerTyrThrPheSerValTyrIleProSerSerPheProAspAsnAlaThrThrIle 160  
Db 593 CGCAGCTATACCTTTTACGTACATACCTCTCCCTCCCGCAATGCGACACTAATT 652  
QY 161 PheAlaGlnTyrPheHisGlyAlaProSerArgThrLeuValAlaThrProGlnGlyIle 180  
Db 653 TTTCGCCAATGGATGGTGGCCACCAAGAGAGAGCTTGTAGCTTACACCAAGAGGAGAAAT 712  
QY 181 LysThrLeuSerIleGlnGlnPheLeuAlaLeuTyrAspArgMetIlePheLysLysAsn 200  
Db 713 AAAAAGCTGAGCATGAGAAGTTTGGCTTATACGACCGCATGATCTTAAAAAAAT 772  
QY 201 IleAlaHisAspLysValGlnLysLysAspLysAspLysLysIleThrTyrValAlaGly 220  
Db 773 ATGCCCATATTAAGTTGAAAAAAGATTAAGACGGAATAATTACTTATGAGCGGGA 832  
QY 221 LysProAsnGlyTyrPheLysValGlnGlnGlyLysTyrProThrLeuAlaIlePheSer 240  
Db 833 AAGCAAAATGCGTGGAGGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 892  
QY 241 PheThrIleLeuAlaIleAspAsnSerAspArgGlnTyrLeuThrThrAspLysAlaAsp 260  
Db 892

Db 893 AAAGGTATTTTACATCAAGCAAACTCCGACCGGAGTGGCTTACGCAAAAGCCGAC 952  
QY 261 ArgAsnAsnAlaAsnProGlnAsnSerGlnValMetLysProTyrSerSerGlyTyrLys 280  
Db 953 CGTACACATGCGCAATCCCGAAATAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 1012  
QY 281 ThrSerThrIleAlaTyrLysMetProPheAlaGlnPheProLysAspCysTyrIleThr 300  
Db 1013 ACTCAACCATGCTCTTAAAAATCCCTTTGGCCAGTTCCCTAAAGATTCGATGATTACT 1072  
QY 301 PheAspValAlaIleAspTyrThrLysTyrGlyLysGlnAlaAsnThrIleLeuLysPro 320  
Db 1073 TTGATGTCGCCATAGACTGAGCAAAATATGGAAGAAAGGCAATATTTTAAACCC 1132  
QY 321 GlyLysLeuAspValMetMetThrTyrThrLysAsnLysLysProGlnLysAlaHisIle 340  
Db 1133 GGTAAAGCTGGATGGATGATGATCTTAACCAAGATTAAGAAACCAAAAGCCCATATC 1192  
QY 341 ValAsnGlnGlnIleLeuIleGlyArgAsnAspAspAspGlyTyrTyrPheLysPhe 360  
Db 1193 GTAAACAGCAGAGAAATCCGATGAGACGTRACATACGATGCTTACTTCAAAATTT 1252  
QY 361 GlyIleTyrArgValGlyAsnSerThrValProValThrTyrAsnLeuSerGlyTyrSer 380  
Db 1253 GGAATTTACAGGCTCGGTAAACAGCAGCAGCGTCCGGTACTTATACCTGAGGGGTACAGC 1312  
QY 381 GluThrAlaArg 384  
Db 1313 GAACCTGCCAGA 1324

## RESULT 2

PCT-US92-09124-1

Sequence 1, Application PC/TUS9209124

## GENERAL INFORMATION:

APPLICANT: Massachusetts Institute of Technology

TITLE OF INVENTION: The Heparinase Gene from Flavobacterium

NUMBER OF SEQUENCES: 2

## CORRESPONDENCE ADDRESS:

ADDRESS: Kipatrick &amp; Cody

STREET: 1100 Peachtree Street, Suite 2800

CITY: Atlanta

STATE: Georgia

COUNTRY: US

ZIP: 30309-4530

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US92/09124

FILING DATE: 09921022

## CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Padst, Pate L.

REGISTRATION NUMBER: 31,284

REFERENCE/DOCKET NUMBER: MIT5546

TELEPHONE: 404-815-6508

TELEFAX: 404-815-6555

INFORMATION FOR SEQ ID NO:

SEQUENCE CHARACTERISTICS:

LENGTH: 1379 base pairs

TYPE: NUCLEIC ACID

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Flavobacterium heparinum

PCT-US92-09124-1

OM protein - protein search, using sw model

Run on: November 20, 2002, 14:41:43 : Search time 43 Seconds

(without alignments)  
858.503 Million cell updates/sec

Title: US-09-802-285-2

Sequence: 1 MKKQILYLVVLOQLFLCSAV.....VGNSTVPVTYNI.SGSETAR 384

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR.73:\*

1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2043	100.0	384	2 A47479	heparin lyase (EC
2	111	5.4	398	2 C86650	hypothetical prote
3	107	5.2	409	2 E69847	hypothetical prote
4	105	5.1	363	2 AE0616	cuter membrane pro
5	105	5.1	725	2 D81976	probable ferric si
6	105	5.1	725	2 H81030	ionb-dependent rec
7	103	5.0	1452	1 S17669	protein-tyrosine-p
8	102.5	5.0	545	2 A38447	coligopeptide ABC t
9	102.5	5.0	1243	2 S60138	sex factor aggrega
10	102	5.0	572	2 AD1209	N-acetylmuramyl-L
11	101.5	5.0	2044	2 AB1180	probable peptidogl
12	100.5	4.9	381	2 S45109	beta-lactamase (EC
13	99.5	4.9	612	2 G83307	hypothetical prote
14	99	4.8	363	2 S43159	outer membrane por
15	98.5	4.8	416	2 A97093	probable membrane
16	98.5	4.8	2523	2 T18477	hypothetical prote
17	98	4.8	515	2 T09272	probable tail comp
18	98	4.8	546	2 D86811	hypothetical prote
19	98	4.8	1452	1 S17670	mannosyl-9-glycopro
20	97.5	4.8	537	2 A35400	protein-tyrosine-p
21	97.5	4.8	1873	2 T30944	surface protein T6
22	97	4.7	790	2 T49542	surfact protein pr
23	97	4.7	1628	2 E90538	xylan 1,4-beta-xy
24	96.5	4.7	1009	2 C64483	hypothetical prote
25	95.5	4.7	381	2 S38196	hypothetical prote
26	95.5	4.7	953	2 S55156	beta-lactamase (EC
27	95	4.7	684	2 S41788	probable membrane
28	95	4.7	1077	2 D71660	endo-1,4-beta-xyla
29	95	4.7	1092	2 S42798	hypothetical prote
					fibronectin-bindin

30	94.5	4.6	381	2 S08296	beta-lactamase (EC
31	94.5	4.6	421	2 G97319	hypothetical prote
32	94.5	4.6	657	2 E81119	tail fibre protein
33	94.5	4.6	718	1 ALBSMX	cyclomalodextrin
34	94	4.6	761	1 B64506	DNA topoisomerase
35	94	4.6	823	2 S14055	nucleoskeletal-lik
36	93.5	4.6	610	2 T22909	hypothetical prote
37	93.5	4.6	710	2 S22673	ferritoxamine recep
38	93	4.6	770	2 A11769	autolysin, amidase
39	93	4.6	1750	2 H64430	ribonucleoside-tri
40	92.5	4.5	315	2 AC0392	probable exported
41	92.5	4.5	422	2 F70018	multiple sugar-din
42	92.5	4.5	1520	2 T44231	hypothetical prote
43	92	4.5	500	2 B43577	perfringolysin O p
44	92	4.5	515	2 T09203	probable tail comp
45	92	4.5	652	2 G95177	hypothetical prote

ALIGNMENTS

RESULT 1

heparin lyase (EC 4.2.2.7) I precursor - Flavobacterium heparinum  
C:Species: Flavobacterium heparinum  
C:Date: 21-Jan-1994 #sequence\_revision 18-Nov-1994 #text\_change 20-Sep-1999  
C:Accession: A47479; S65553  
R:Sasisekharan, R.; Bulmer, M.; Moremen, K.W.; Cooney, C.L.; Langer, R.  
Proc. Natl. Acad. Sci. U.S.A. 90, 3660-3664, 1993  
A:Title: Cloning and expression of heparinase I gene from Flavobacterium heparinum.  
A:Reference number: A47479; MUID:93234557; PMID:8475114  
A:Accession: A47479  
A:Status: preliminary  
A:Molecule type: DNA; protein  
A:Residues: 1-384 <SAs>  
A>Note: sequence extracted from NCBI backbone (NCBI:129766, NCBI:129769)  
R:Ernst, S.; Venkataraman, G.; Winkler, S.; Godavarti, R.; Langer, R.; Cooney, C.L.;  
Biochem. J. 315, 589-597, 1996  
A:Title: Expression in Escherichia coli, purification and characterization of heparin  
A:Reference number: S65553; MUID:96202015; PMID:8615834  
A:Accession: S65553  
A:Molecule type: protein  
A:Residues: 22-27 <ERNS>  
C:Superfamily: Flavobacterium heparinum heparin lyase I  
C:Keywords: carbon-oxygen lyase

Query Match 100.0% Score 2043; DB 2; Length 384;  
Best Local Similarity 100.0%; Pred. No. 9.2e-152;  
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MKKQILYLVVLOQLFLCSAVAAQOKSGNIPYRVNQADSAKOKAIIDNKVAVGINKPYA 60	
DB	1	MKKQILYLVVLOQLFLCSAVAAQOKSGNIPYRVNQADSAKOKAIIDNKVAVGINKPYA 60	
QY	61	LOYDCKLRNKPSTRFELKADNSLEGYAGETGRFELSYVATNDKFFPSVYON 120	
DB	61	LOYDCKLRNKPSTRFELKADNSLEGYAGETGRFELSYVATNDKFFPSVYON 120	
QY	121	AKKLTIVYHKGICOGSSRSYTSVIPSFPDNATIFQMGHGPRTLVATPEGFI 180	
DB	121	AKKLTIVYHKGICOGSSRSYTSVIPSFPDNATIFQMGHGPRTLVATPEGFI 180	
QY	181	KTLSTEEFLATYDRIAFKKNIAHDKVEKKDKGKITYVAGKPNKGKVEGQGYPTLAFGRS 240	
DB	181	KTLSTEEFLATYDRIAFKKNIAHDKVEKKDKGKITYVAGKPNKGKVEGQGYPTLAFGRS 240	
QY	241	KGYFIKANSDDQWLTDKADRNANPENSEVAKPYSSSEYKSTIAYKMPFAOPPKDCWIT 300	
DB	241	KGYFIKANSDDQWLTDKADRNANPENSEVAKPYSSSEYKSTIAYKMPFAOPPKDCWIT 300	
QY	301	FVVAIDWTKYGEANTILKPGKLDVAMTYYTKKKKQKAKHIVNQGHILGRNDDGYRFE 360	
DB	301	FVVAIDWTKYGEANTILKPGKLDVAMTYYTKKKKQKAKHIVNQGHILGRNDDGYRFE 360	

QY 361 GIVRGSTVPYVYNLSGSETAR 384  
 Db 361 GIVRGSTVPYVYNLSGSETAR 384

## RESULT 2

C86650

hypothetical protein rpgD (imported) - Lactococcus lactis subsp. lactis (strain IL1403)

C:Species: Lactococcus lactis subsp. lactis

C:Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 24-Aug-2001

C:Accession: C86650

R:Bolotin, A.; Winkler, P.; Mueger, S.; Jallou, O.; Malarme, K.; Weissenbach, J.; Ehrlich

Genome Res. 11, 731-753, 2001

A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ss

A:Reference number: AB6625; MUID:21235186; PMID:11337471

A:Accession: C86650

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-398 &lt;STO&gt;

A:Cross-references: GB:AE005176; PID:g12723058; PIDN:AAK04301.1; GSPDB:GN00146

A:Experimental source: strain IL1403

C:Genetics:

A:Gene: rpgD

C:Superfamily: Streptococcus mutans ABC transporter rpgDc: ATP-binding cassette homology

## Query Match

Best Local Similarity 5.4%; Score 111; DB 2; Length 398;

Matches 82; Conservative 43; Mismatches 125; Indels 108; Gaps 20;

QY 41 KOKAIDNKKVVA-----VGINKPYALOYDOKLRNKGKPSYRFELEKADNSIEGAAETK 95

Db 84 KGVNIYNGKLVSTIEAGVGNP-----ELTGKENV--YLNQAML 120

QY 96 G--RTELSSYATNTDFKPKPPSYQNAOKLKYHYHGKICQEGSSRSYTFYSPSSF 153

Db 121 GFSREISDMYDEIVDAEL--EEFNM-QRLK--NYSQGM-----QVRLAFSAIKNA-- 161

QY 154 PDNATTFQOMHGAPSTVLATPEGEIKTISIEFEALYDPMFKKNIANDYKCKDKG 213

Db 168 -----RGDV--LVLEVLAVGDESFQK--CNDYFPRKKS 200

QY 214 KITTY-----AGK-----PNGKRVGOGGPTLAFGSKGYF--IKANS 251

Db 201 KITLIVTHMGAAKKYCNKAILLENGLIKSGVDVNDVANSIDNLQEVYARSDLEIDV 260

QY 252 -RQWLFDKADRNANPE-----NSEVMKPYSEYKSTIAY-KMPFAQFPKDCWTFDVAI 305

Db 261 KEEMI-ENLEVNFINQOKIPNDEVKFEVSVDYKKQETVAFSELTVDNRINWYNLSL 319

QY 306 DMTRYCKEANTILPKGLD-----VMMYTKKKKPKKAHYNOEILGRND--DD 354

Db 320 DYLTGKGNHRAIYTCKLDOVNNMLKLOYSIRKRMDEMAFDEKYLINRDISD 377

QY 320 DYLTGKGNHRAIYTCKLDOVNNMLKLOYSIRKRMDEMAFDEKYLINRDISD 377

Db 320 DYLTGKGNHRAIYTCKLDOVNNMLKLOYSIRKRMDEMAFDEKYLINRDISD 377

QY 320 DYLTGKGNHRAIYTCKLDOVNNMLKLOYSIRKRMDEMAFDEKYLINRDISD 377

Db 320 DYLTGKGNHRAIYTCKLDOVNNMLKLOYSIRKRMDEMAFDEKYLINRDISD 377

QY 320 DYLTGKGNHRAIYTCKLDOVNNMLKLOYSIRKRMDEMAFDEKYLINRDISD 377

Db 320 DYLTGKGNHRAIYTCKLDOVNNMLKLOYSIRKRMDEMAFDEKYLINRDISD 377

QY 320 DYLTGKGNHRAIYTCKLDOVNNMLKLOYSIRKRMDEMAFDEKYLINRDISD 377

Db 320 DYLTGKGNHRAIYTCKLDOVNNMLKLOYSIRKRMDEMAFDEKYLINRDISD 377

QY 320 DYLTGKGNHRAIYTCKLDOVNNMLKLOYSIRKRMDEMAFDEKYLINRDISD 377

Db 320 DYLTGKGNHRAIYTCKLDOVNNMLKLOYSIRKRMDEMAFDEKYLINRDISD 377

QY 320 DYLTGKGNHRAIYTCKLDOVNNMLKLOYSIRKRMDEMAFDEKYLINRDISD 377

Db 320 DYLTGKGNHRAIYTCKLDOVNNMLKLOYSIRKRMDEMAFDEKYLINRDISD 377

QY 320 DYLTGKGNHRAIYTCKLDOVNNMLKLOYSIRKRMDEMAFDEKYLINRDISD 377

Db 320 DYLTGKGNHRAIYTCKLDOVNNMLKLOYSIRKRMDEMAFDEKYLINRDISD 377

QY 320 DYLTGKGNHRAIYTCKLDOVNNMLKLOYSIRKRMDEMAFDEKYLINRDISD 377

Db 320 DYLTGKGNHRAIYTCKLDOVNNMLKLOYSIRKRMDEMAFDEKYLINRDISD 377

QY 320 DYLTGKGNHRAIYTCKLDOVNNMLKLOYSIRKRMDEMAFDEKYLINRDISD 377

Db 320 DYLTGKGNHRAIYTCKLDOVNNMLKLOYSIRKRMDEMAFDEKYLINRDISD 377

QY 320 DYLTGKGNHRAIYTCKLDOVNNMLKLOYSIRKRMDEMAFDEKYLINRDISD 377

Db 320 DYLTGKGNHRAIYTCKLDOVNNMLKLOYSIRKRMDEMAFDEKYLINRDISD 377

QY 320 DYLTGKGNHRAIYTCKLDOVNNMLKLOYSIRKRMDEMAFDEKYLINRDISD 377

Db 320 DYLTGKGNHRAIYTCKLDOVNNMLKLOYSIRKRMDEMAFDEKYLINRDISD 377

QY 320 DYLTGKGNHRAIYTCKLDOVNNMLKLOYSIRKRMDEMAFDEKYLINRDISD 377

Db 320 DYLTGKGNHRAIYTCKLDOVNNMLKLOYSIRKRMDEMAFDEKYLINRDISD 377

QY 320 DYLTGKGNHRAIYTCKLDOVNNMLKLOYSIRKRMDEMAFDEKYLINRDISD 377

Db 320 DYLTGKGNHRAIYTCKLDOVNNMLKLOYSIRKRMDEMAFDEKYLINRDISD 377

QY 320 DYLTGKGNHRAIYTCKLDOVNNMLKLOYSIRKRMDEMAFDEKYLINRDISD 377

Db 320 DYLTGKGNHRAIYTCKLDOVNNMLKLOYSIRKRMDEMAFDEKYLINRDISD 377

QY 320 DYLTGKGNHRAIYTCKLDOVNNMLKLOYSIRKRMDEMAFDEKYLINRDISD 377

Db 320 DYLTGKGNHRAIYTCKLDOVNNMLKLOYSIRKRMDEMAFDEKYLINRDISD 377

QY 320 DYLTGKGNHRAIYTCKLDOVNNMLKLOYSIRKRMDEMAFDEKYLINRDISD 377

Db 320 DYLTGKGNHRAIYTCKLDOVNNMLKLOYSIRKRMDEMAFDEKYLINRDISD 377

QY 320 DYLTGKGNHRAIYTCKLDOVNNMLKLOYSIRKRMDEMAFDEKYLINRDISD 377

Db 320 DYLTGKGNHRAIYTCKLDOVNNMLKLOYSIRKRMDEMAFDEKYLINRDISD 377

QY 320 DYLTGKGNHRAIYTCKLDOVNNMLKLOYSIRKRMDEMAFDEKYLINRDISD 377

T.: Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida  
 A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.  
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis  
 A:Reference number: AB6580; MUID:98044033; PMID:9384377

A:Accession: E69847

A:Status: preliminary

A:Molecule type: nucleic acid sequence not shown; translation not shown

A:Residues: 1-409 &lt;KUN&gt;

A:Cross-references: GB:Z99110; GB:AL009126; MUID:92633472; PIDN:CA13048.1; PID:el183;

A:Experimental source: strain 168

C:Genetics:

A:Gene: yjcm

C:Superfamily: Bacillus subtilis/hypothetical protein yjcm

## Query Match

Best Local Similarity 2.2%; Score 107; DB 2; Length 409;

Matches 85; Conservative 56; Mismatches 155; Indels 126; Gaps 15;

QY 1 MKKQILYKLV-----QQLPLCSAVAQKSGNIPYRVNVAADSAROKAIIIDNKW 50

Db 1 MKKQILYKLV-----QQLPLCSAVAQKSGNIPYRVNVAADSAROKAIIIDNKW 50

QY 51 VAVGINKPYALOYDDKLRNKGKPSYRFELEKADNSIEGAAETKRTLSYATNTDF 110

Db 57 YDFEHHTPLAHPRKSL-----TDSQEGV-----LTSALSKVDK 91

QY 111 KRPFSYQNAOKLKYHYHGKICQEGSSRSYTFYSPSSFPDN----- 156

Db 92 QKHETLY-DIDVKSSEKNNKILYEAITRNVPFVEKAYTTLGDNKILEPKTTSNKO 150

QY 157 -----ATTFAQMHGAPSTVLATPEGEIKTISIEFEALYDPMFKKNIANDYKCKDKG 185

Db 151 SSOSTNTSNLSFSQVKNSSNQLNVIPLSNSSQOLNTQLNNSSTLYQENDEEYSSDVKL 210

QY 186 EEFLLAYDMIEFKKNAHDKVEKKDKITTYAAGPNGKRYVQGGPTLAFPSGYY-- 243

Db 211 DEFLKTKQEV--KNDKREKSEDKPVSNFNAANSNOLTV--ATSVKGS-GYAA 263

QY 244 -----FYKANSRQMLDTPKADRNANPENSEVMKPYSEYKSTIAYKMPFAQFPKDC 297

Db 264 MKYAEYALKPKNNNYKDKDCTNVSQALRGMRPFYEMKPYDA----- 311

QY 298 WTFDVAIDMTYKKEANTILKPKGL-----DVMYTKKKKPKKAHYNOEILGRND 354

Db 312 -----VWNAQAFRSYILKAGIKMKIVSDTYSNVKLGIDYHYDHNK--IGLRAD 360

QY 355 GX 356

Db 361 GW 362

QY 361 GW 362

Db 361 GW 362

QY 361 GW 362

Db 361 GW 362

QY 361 GW 362

Db 361 GW 362

QY 361 GW 362

Db 361 GW 362

QY 361 GW 362

Db 361 GW 362

QY 361 GW 362

Db 361 GW 362

QY 361 GW 362

Db 361 GW 362

QY 361 GW 362

Db 361 GW 362

QY 361 GW 362

Db 361 GW 362

QY 361 GW 362

Db 361 GW 362

QY 361 GW 362

Db 361 GW 362

QY 361 GW 362

Db 361 GW 362

QY 361 GW 362

Db 361 GW 362

QY 361 GW 362

Db 361 GW 362

QY 361 GW 362

Db 361 GW 362

QY 361 GW 362

Db 361 GW 362

QY 361 GW 362

Db 361 GW 362

QY 361 GW 362

Db 361 GW 362

QY 361 GW 362

Db 361 GW 362

QY 361 GW 362

Db 361 GW 362

QY 361 GW 362

Db 361 GW 362

QY 361 GW 362

Db 361 GW 362

QY 361 GW 362

Db 361 GW 362